

WO 00/58342

1

PCT/FI00/00249

## SEQUENCE LISTING

&lt;110&gt; Valtion teknillinen tutkimuskeskus

&lt;120&gt; Process for partitioning of molecules

&lt;130&gt; 31805

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 42

&lt;170&gt; PatentIn Ver. 2.2

&lt;210&gt; 1

&lt;211&gt; 428

&lt;212&gt; DNA

&lt;213&gt; Trichoderma reesei

&lt;220&gt;

&lt;221&gt; intron

&lt;222&gt; (167)...(236)

&lt;220&gt;

&lt;221&gt; intron

&lt;222&gt; (323)...(386)

&lt;220&gt;

&lt;223&gt; Coding sequence of hfbl

&lt;400&gt; 1

atgaagtct tcgccatcg cgcgtctttt gcccggctg ccgttgccca gcctctcgag 60  
gaccgcagca acggcaacgg caatgtttgc cttccggcc tcttcagcaa cccccagtgc 120  
tgtgccaccc aagtccctgg cctcatcgcc ctgtactgca aagtccgtaa gttgagccat 180  
aacataagaa tccctttggc gaaatatgc cttctcactc cttaaccctt gaacagcrtc 240  
ccagaacgtt tacgacggca cggacttccg caacgtctgc gccaaaaccg gcgcggcagcc 300  
tctctgtgc gtggccccc ttgttaagttt atgccccagg tcaagctcca gtcttggca 360  
aacccattct gacaccaga ctgcaggccg gccaggctct tctgtgccag accgcgtcg 420  
gtgcttga 428

&lt;210&gt; 2

&lt;211&gt; 78

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: PCR 5' primer

&lt;400&gt; 2

tccggcacta cgtgccagta tagcaacgac tactactcg aatgccttgt tccgcgtggc 60  
tcttagttctg gaaccgca 78

<210> 3  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 3  
tcgtacggat cctcaaggcac cgacggcggt 30

<210> 4  
<211> 63  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 4  
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcagag caacggcaac 60  
ggc 63

<210> 5  
<211> 2211  
<212> DNA  
<213> *Trichoderma reesei*

<220>  
<221> promoter  
<222> (1)..(2211)  
<223> *cbh1* promoter sequence

<400> 5  
gaattctcac ggtgaatgt a ggccttttgt agggtaggaa ttgtcactca agcaccggca 60  
acctccatta cgcctcccc atagagttcc caatcagtga gtcatggcac tggctcaaa 120  
tagattgggg agaagttgac ttccggccag agctgaaggt cgccacaaccg catgatata 180  
ggtcggcaac ggcaaaaaag cacgtggctc accgaaaagc aagatgttg cgtatctaaca 240  
tccaggaacc tggatacatc catcatcagc cacgaccact ttgatctgtt ggtaaactcg 300  
tattcgccct aaaccqaagt gctgtggtaaa tctacacgtt ggccttcc ggtataactgc 360  
gtgtgtcttc tctaggtgca ttctttctt cctctagtgt tgaattgttt gtgttggag 420  
tccgagctgt aactacctct gaatctctgg agaatgttgg actaacgact accgtgcacc 480  
tgcatacatgt atataatagt gatcctgaga aggggggtt ggagcaatgt gggactttga 540  
tggtcatcaa acaaagaacg aagacgcctc ttttgc当地 gctacggta 600  
agaactggat acttgttgc当地 ttttgc当地 caacaagagg ccagagacaa 660  
tctattcaaa caccaagctt gctcttttga gctacaagaa cctgtggggat atatatctag 720  
agtgtgaag tcggtaatcc cgctgtatag taatacggat cgcatctaa tactccgaag 780  
ctgctgc当地 cccggagaat cgagatgtgc tggaaagctt ctggcggcg gctaaattag 840  
catgaaaggc tatgagaaaat tctggagacg gcttgc当地 tcatggcggtt ccattcttc当地 900  
acaagcaag cgttccgtcg cagtagcagg cactcatcc cgaaaaactt cgagattcc 960  
taatgtgc当地 tggaaaccgg当地 ataataat aggcaataca ttgagttgcc tggacgggtt 1020  
caatgc当地 gtaactggat tggacataac tggctcgatccc当地 cttcaaccctt当地 1080  
ggcggttcc当地 tgattcagcg tacccgtaca agtgc当地 atctatataacc cagactgacc 1140  
ggacgtgttt tgcccttcat ttggagaaaat aatgtc当地 tggatgttgc当地 1200  
tgaccgactg gggctgtcg aagcccgaaat gtggatgtt tattccgact ctgctcgat 1260  
aggcatgtt gtaatctgtg tcggcagga caccgc当地 cgg当地 aggttccacgg caaggaaac 1320

caccgataggc agtgtctagt agcaacctgt aaagccgcaa tgcagcatca ctggaaaata 1380  
caaaccaatg gctaaaagta cataagttaa tgccctaaaga agtcatatac cagccgctaa 1440  
taattgtaca atcaagtggc taaacgtacc gtaatttgcc aacgcgttgtt ggggttgcag 1500  
aagcaacggc aaagcccact tcccacgtt gtttcttcac tcagtc当地 ctcagcttgt 1560  
gatcccccaa ttgggtcgct tgggttcc ggtgaagtga aagaagacag aggtaaagaat 1620  
gtctgactcg gagcgttttgc catacaacca agggcagtga tggaaagacag tgaatgttg 1680  
acattcaagg agtatttagc caggatgtc tgagtgtatc gtgtaaaggag gtttctgc 1740  
cgatacgac aatactgtat agtacttct gatgaagtgg tccatatgaa aatgtaaagtc 1800  
ggcactgaac aggcaaaaaga ttgagttgaa actgcctaaag atctcgccctc ctcgggcttc 1860  
ggcttgggtt gtacatgtt tgctccggg caaatgaaa agtggtagg atcgacarac 1920  
tgctgcctt accaaggcgc tgagggtatg tgataggaa atgttcaggg gccactgcat 1980  
gtttcgaat agaaaagaaaa gcttagccaa gaacaatagc cgataaaagat agcctcatta 2040  
aacgaaatga gctagtaggc aaagtcaagc aatgtgtata tataaaggtt cgaggtccgt 2100  
gcctccctca tgctctcccc atctactcat caactcagat cctccaggag acttgtacac 2160  
catctttga ggcacagaaaa cccaatagc aaccgcggac tgcgcatcat g' 2211

<210> 6  
<211> 1588  
<212> DNA  
<213> *Trichoderma reesei*

<220>  
<223> *T. reesei* egli cDNA

<400> 6  
ccccctatc ttagtccttc ttgttgccttcc aaaatggcgc cctcagttac actgccgttg 60  
accacggcca tcctggccat tgcggggtc gtcggccccc agcaacccggg taccagccccc 120  
cccgagggtcc atcccaagtt gacaacctac aagtgtacaa agtccggggg gtgcgtggcc 180  
caggacacctt cgggtgtct tgacttggaa taccgttggaa tgcacgcgc aaactacaac 240  
tcgtgcaccc tcaacggcggg cgtcaacacc acgtctgtcc ctgacggggc gaccgtggc 300  
agaactgtct tcatcgaggg cgtcgactac gcccgttggg gcgtcacac ctcgggcaggc 360  
agcctcacca tgaaccagta catggccaggc agtctgggg gctacagcag cgtctctcc 420  
cggctgtatc tcttggactc tgacggtagt tacgtgtatc tgaagctaa cggccaggag 480  
ctgagcttcg acgtcgacact ctctgtctg ccgtgtgggg agaacggcgc gctctacctg 540  
tctcagatgg acgagaacagg gggcccaac cagttataaca cggccgtgc caactacggg 600  
agcggctact gcgatgtca gtgcggccgc cagacatggaa ggaacggcac cctcaacact 660  
agccaccagg gtttctgtc caacggatg gatatcctgg agggcaactc gagggcaaat 720  
gccttggccctt ctcacttctg cacggccacg gcctgtcgact ctgcccgggtt cggcttcaac 780  
ccctatggca ggggtacaaa aagctactac ggccccgggg ataccgttga caccttcaag 840  
accttcacca tcatcacccca gttcaacacg gacaacggct cggccctggg caaccttgg 900  
agcatcaccc gcaagtatcca gcaaaacggc gtgcacatcc ccagcgccca gcccggcggc 960  
gacaccatct cgtccctggcc gtcggccctca gcctacggcg gcctcgccac catggcaag 1020  
gccttggccat gggcatgtt gtcgtgttc agcatgttggaa acgacaacacg ccagtacatg 1080  
aactggctcg acagccggaa cggccggccccc tgcggccggc acggaggccaa cccatccaac 1140  
atccctggcca acaaccccaa cacgcacgtc gtcttctccaa acatccgttgggg gggagac 1200  
gggtctacta tgaactcgac tgcggccccc ccccccgtt cgtccagcac gacgtttcg 1260  
actacacggg ggagctcgac gacttcgac agcccgagct gcacgcagac tcaactggggg 1320  
cagtgccgtt gcatggggta cagccgggtgc aagacgtgc acgtccggccac tacgtggccag 1380  
tatagcaacg actactactc gcaatgcctt tagacgttgtt acttgcctt ggtctgtcca 1440  
gacggggca cgatagaatg cggccacgc gggagctgtt agacattggg cttatatat 1500  
aagacatgtt atgttgcatac tacattagca aatgacaaac aatgaaaaaa gaacttatca 1560  
agcaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa 1588

<210> 7  
<211> 745  
<212> DNA  
<213> *Trichoderma reesei*  
  
<220>  
<221> terminator  
<222> (1)..(745)  
<223> *T. reesei cbh1* terminator

<400> 7  
ggaccttacc agtctcacta cggccagtgc ggcggatttgc gctacagcgg ccccacggtc 60  
tgcggccagcg gcacaacttg ccagggtcctg aacccttact acttcagtg cctgtaaagc 120  
tccgtgcga agcctgacgc accggtagat tcttggtgag cccgtatcat gacggcggcg 180  
ggagctacat ggccccgggt gatttatttt ttttgtatct acttctgacc ctttcaaat 240  
atacggtaaa ctcatctttc actggagatg cggcctgctt ggtattgcga tgggtcagc 300  
ttggcaaatt gtggctttcg aaaacacaaa acgattcctt agtagccatg cattttaaaga 360  
taacggaaata gaagaaagag gaaattaaaa aaaaaaaaaa aacaaacatc cggttcataa 420  
cccgtagaat cggcgtctt cgtgtatccc agtaccacgt caaaggtatt catgatcggt 480  
caatgttcat attgttccgc cagttatggct ccaccccat ctccgcgaat ctccctttct 540  
cgaacgcggt agtggctgct gccaattggt aatgaccata gggagacaaa cagcataata 600  
gcaacagtgg aaatttagtgg cgcaataatt gagaacacag tgagaccata gctggcggcc 660  
tggaaagcac tggggagac caacttgtcc gttgcgaggc caacttgcatt tgctgtcaag 720  
acgatgacaa cgtagccgag gaccc 745

<210> 8  
<211> 10  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: annealed primer

<400> 8  
taaccgcgggt

10

<210> 9  
<211> 16  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: annealed primer

<400> 9  
ctagaccgcg gttaat

16

<210> 10  
<211> 1232  
<212> DNA  
<213> *Trichoderma reesei*  
  
<220>  
<221> promoter  
<222> (1)..(1232)

<223> *T. reesei gpd1* promotor

<400> 10

gtcgacacga tatacaggcg cggtctatgtaa taatgtatgtat cgagcatgac ttgtatgtgt 60  
atgtgacaat attgactgcg aggaaccatc aggtgtgtat ggatggaaatc atttctgttaac 120  
caccagggtg catgtatcat aaggattctc ctcagctcac caacaacgaa cgatggccat 180  
gttagtgaag gcaccgtat ggcaagatag aaccactatt gcatctgcgc ttcccacgca 240  
cagtacgtca agtaacgtca aagccgcct cccgttaacct cggccgttgt tgctccc 300  
gattgcctca atcacatgt acctacctat gcattatggg cggcctcaac ccacaaaaaaa 360  
agattgagag ctacccatac tcaatatggc cagcacctct tcggcgatac atactcgcca 420  
ccccagccgg cgcgattgtg tgtaacttaggt aggctcgatc tataccagca ggagaggtgc 480  
tgcttggcaa tcgtgtctcg ctgttaggtt gtactttgtat ggtaacttgcgat 540  
gcagttgtca aggtacctag ggagggttc aacgaggccct gcttcaatgcgatcc 600  
taggatggcg gctgggggg cagaagctgg gaactcgcca acagtcataat gtaatagctc 660  
aagttgtatc taccgttttgc ctagatttttgc tgcgagaagc agcatgaatg tcgctcatcc 720  
gatggcgcat caccgttgcg tcaaaaaacga ccaagctaaatc aactaaggtaatcc 780  
ccactatctc aggttaaccatc gtactaccatc ctaccctacc tgccgtgcctt accctgtttt 840  
gtgttaatctt ttcacccctcc ctcttcataatc ttcttttttttcc ttttttttttt 900  
cttccttcctc ttcttcataatc taacccattcc taacaacatc gacattttctt cctaatacc 960  
agcctcgccaa atccctatgtt tgatgtacg tacgtactac aatcatcacc acgatcgcc 1020  
gccccgacat gccccgttctg ttccgtctgcc cctcccttcataatc ctccgtccct tgacgagatc 1080  
gccccggccat gacttcctgc cgtccaccaat ttttttccctt atttacccctt cctcccttcataatc 1140  
tccctctctgt ttcttcataatc caaacaacca ccacccaaat ctctttggaa gctcacgact 1200  
caccgcgatc aatttcgcaga tacaatctca ga 1232

<210> 11

<211> 1129

<212> DNA

<213> *Trichoderma reesei*

<220>

<221> terminator

<222> (1)..(1129)

<223> *T. reesei gpd1* terminator

<400> 11

ggatccccgag cattgtctat gaatgcacaaac aaaaatagta aataaatatgtt aattctggcc 60  
atgacgaata gagccaaatct gctccacttgc actatcttgc gactgtatcg tatgtcgaaac 120  
ccttgactgc ccatttcaac aattgtaaaatggaaatgtt acaaggatgtt tctcacatgtt 180  
gcgtgcgagc ccgtttgtac gtatgttgc gaaaggcgttgc ccacatgttgc tctcacatgtt 240  
cttggcttac gatcatgtttt gcatgttgc gtaagaatc acagatgttgc gattatctcc 300  
atcgcttcataatc tgatgttgc ctcagacaaatc acatgttgc gaaaggatgttgc atcgcatgtt 360  
aggtcgttgc caatcatgttgc ctggacttgc gttatgttgc taagtcataatc tacccttcataatc 420  
agaaggatgttgc gtagggatgttgc tccggatgttgc accatgttgc ctgcataatc aagaaacccc 480  
ccaaaaggatgttgc tccggatgttgc accatgttgc acatgttgc aatgttgc aaaaaggatgttgc 540  
agaatgttgc tccggatgttgc aatgttgc tccggatgttgc tccggatgttgc aatgttgc aaaaaggatgttgc 600  
aacaatccatgttgc tccggatgttgc aatgttgc tccggatgttgc tccggatgttgc aatgttgc aaaaaggatgttgc 660  
tgaggtgacc tgcgttttttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 720  
ccggccctcg cctcttcataatc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 780  
gaaggcggttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 840  
atcaaaatccatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 900  
accggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 960  
gttccatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 1020  
aggagagatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 1080  
gaatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc tccggatgttgc 1129

<210> 12  
<211> 5733  
<212> DNA  
<213> *Aspergillus nidulans*

<220>  
<223> (1-5733) Sequence of plasmid pAN52-1

<220>  
<221> promoter  
<222> (1)..(2129)  
<223> *A. nidulans gpdA* promoter

<220>  
<221> gene  
<222> (2130)..(2304)  
<223> *A. nidulans gpdA* gene

<220>  
<221> terminator  
<222> (2305)..(3071)  
<223> *A. nidulans trpC* terminator

<220>  
<221> misc\_feature  
<222> (3072)..(5726)  
<223> pUC18 from SalI to EcoRI

<400> 12  
caattccctt gtatctctac acacaggctc aaatcaataa gaagaacgggt tcgtcttttt 60  
cgtttatatac ttgcategtc ccaaagctat tggcgggata ttctgtttgc agttggctga 120  
cttgaagtaa tctctgcaga tctttcgaca ctgaaatacg tcgagcctgc tccgcttgg 180  
agcggcgagg agcctcgccc tgcacaact accaacatgg agtacgatgg gggccagttc 240  
cgccagctca ttaagagccca gtcatgggc gttggcatga tggccgtcat gcatctgtac 300  
ttcaagtaca ccaacgtctc tctgtatcccg tcgatcatcc gctgaaggcg ctttcgaatc 360  
tggttaagat ccacgtcttc gggaaaggcccg cgactgggtg cctccagcgt ccctttaagg 420  
ctgccaacag ctttctcagc caggcccgcc ccaagaccga caaggcctcc ctccagaacg 480  
ccgagaagaa ctggagggggt ggtgtcaagg aggagaatgc tctttattga agtcggagga 540  
cggagcgggtg tcaagaggat attttcgac tctgttattat agataagatg atgaggaatt 600  
ggaggttagca tagcttcatt tgattttgtc ttccaggctg agactctagc ttggaggata 660  
gagggtcctt tggctttcaa tattctcaag tatctcgagt ttgaacttat tccctgtgaa 720  
ccttttatttc accaatgagc attggaaatga acatgaatct gaggactgca atcgccatga 780  
ggttttcgaa atacatccgg atgtcgaagg cttggggcac ctgcgttgg tgaattttaga 840  
acgtggcaact attgatcata cgatagctct gcaaaggcccg ttgcacaatg caagtcaaac 900  
gttgcttagca gttccagggtg gaatgtttagt atgagcatgg tattaaatca ggagatata 960  
catgatctct agttagctca ccacaaaatg cagacggcgt aaccaaaaatg cacacaacac 1020  
aagctgttaag gatttcggca cggctacggg agacggagaa gccaccctca gtggactcga 1080  
gtaccatttatttgc ttgttgcgt agacctaata cagccccctac aacgaccatc 1140  
aaagtctgtat agctaccatg gaggaaagtgg actcaaatcg acttcagcaa catctcctgg 1200  
ataaaacttta agcctaaact atacagaata agatagggtgg agagcttata ccgagctccc 1260  
aaatctgtcc agatcatgtt tgacccggcgt ctggatcttc ctatagaatc atcttttattc 1320  
gttgacccatg ctgattctgg agtgaccggcgt agggtcatga ctggatcttc aaatccggccg 1380  
cctccaccatg ttgtagaaaa atgtgacgaa ctgcgtgacgt ctgtacatgg accgggtgact 1440  
ctttctggca tgcggagaga cggacggacg cagagagaag ggctgagtaa taagccactg 1500  
gccagacacg tctggccggcgt ctggatctgg ataaatccgg gacccggccgc 1560  
ccctccggccg cgaagtggaa aggctgggtgt gcccctcggtt gaccaagaat ctattgcata 1620  
atcggagaat atggagcttc atcgaatcac cggcagtaag cgaaggagaa tggtaagcc 1680  
gggggtgtata gccgtcggcg aaatagcatg ccattaacct agtacagaa gtcacattgc 1740

ttccgatctg gtaaaagatt cacgagatag taccttctcc gaagtaggta gagcggat 1800  
ccggcggtga agctccctaa ttggcccatc cggcatctgt agggcgcca aataatcg 1860  
ctctctctgt ttgccccgtg tatgaaaaccg gaaaggccgc tcaggagctg gccagcggcg 1920  
cagaccggga acacaagctg gcagtcgacc catccggcgc tctgcactcg acctgctgag 1980  
gtccctcagt ccctggtagg cagcttgcc ccgtctgtcc gcccgggtgt tcggcgggt 2040  
tgacaagggtc gttgcgtcag tccaaacattt gttccatat tttctgtc tccccaccag 2100  
ctgctctttt cttttctctt tctttccca tttcagttt attcatcttc ccattcaaga 2160  
accttttattt cccctaagta agtacttgc tacatccata ctccatctt cccatccctt 2220  
attcctttga acctttcagt tcgagcttcc caacttcata gcagcttgac taacagctac 2280  
cccgctttagg cagacatcac catggatcca cttaaacgtta ctgaaatcat caaacagctt 2340  
gacgaatctg gatataagat cggtgggtc gatgtcagct ccggaggatgaa gacaaatgg 2400  
gttccaggatc tcgataagat acgttccatgtt gtcacccatgaa caaagagtgc cttctgtg 2460  
ttaatagctt ccattgtcaac aagaataaaaa cgcgttttcg ggttacccctt tccagatac 2520  
agctcatctg caatgttta atgcatttgac tgcaacccatgaa tttccatctt cgggtccgg 2580  
cgaagagaag aatagcttag cagagatattt ttcattttcg ggagacgaga tcaacagat 2640  
caacggcgta caagagacctt acgagactgaa ggaatccgcctt cttggctcca cggactata 2700  
tatttgtctc taattgtact ttgacatgtt cctttctt actctgtatag cttgactatg 2760  
aaaatccgtt caccagcncc tgggttcgca aagataattt catgtttttt ctttgaactc 2820  
tcaaggcttac aggacacaca ttcatctgtt gtataaaacccat cggaaatccat tccactaag 2880  
atggtataaca atgatcaaccat tgcatttttttgcgtt gtcatttttttgcgtt accaaatacg 2940  
ccggccgaaa ctttttaca actcttcttat gatgttttttgcgtt cccagaatgc acaggtacac 3000  
ttgttttagag gtaatccctt ttcttagaaag ttctctgttgcgtt gtcgttgcgtt gcccactcca 3060  
catctccactt ccgttgcgttgcgtt gcatgcggc ttggacttgcgtt ccgttgcgtt acaacgtcg 3120  
gactggggaaa accctggcgatc taccaactt aatccatcttgcgtt cggccatcc ccccttcg 3180  
agctggcgta atgatcaaga ggcggccaccat gatgttttttgcgtt cccaaatggatg ggcggccgt 3240  
aatggcgat ggcggccatgtt ggcggccatgtt ccccttcgtt atctgttgcgtt tatttcacac 3300  
cgcatatgtt ccgttgcgttgcgtt gcatatgttgcgtt cccatccatcc cccatccatcc 3360  
caccggccaa caccggccatgtt ccgttgcgttgcgtt cccatccatcc cccatccatcc 3420  
agacaagctg tgaccgttcc cggggccatgttgcgtt gatgttgcgtt ggttttaccatgttgcgtt 3480  
aaacggcgta gacgaaagggtt cccatccatcc cccatccatcc tataatgttgcgtt ggttttaccatgttgcgtt 3540  
ataatggttt ctttagacgttcc aggtggcaactt tttccatccatcc tttccatccatcc tttccatccatcc 3600  
tggtttatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 3660  
atgcttcaat aatatttggaa aaggaaagatg atgatgttccatcc tttccatccatcc tttccatccatcc 3720  
atcccccttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 3780  
gtaaaagatg ctgaaatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 3840  
agccgttcaaga tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 3900  
aaagttctgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 3960  
cgccgcataacttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4020  
cttacggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4080  
actccggccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4140  
cacaacatggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4200  
ataccggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4260  
cttacggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4320  
gcggatataatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4380  
gataaaatctgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4440  
ggtaagcccttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4500  
cgaaaatggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4560  
caagtttacttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4620  
taggtgaaga tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4680  
cactggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4740  
ccgtatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4800  
gatcaagatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4860  
aatactgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4920  
cctacatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 4980  
tgtcttaccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 5040  
acgggggggttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 5100  
ctacaggatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 5160  
ccgtatgttccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc tttccatccatcc 5220

tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgctg atttttgtga 5280  
tgctcgta gggggccggag cctatggaaa aacgccagca acgcggcctt ttacggttc 5340  
ctggcccttt gctggccctt tgctcacatg ttctttcctg cgttatcccc tgattctgtg 5400  
gataaccgta ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag 5460  
cgcagcgaatc cagttagcgaa ggaagcgaa gagcgccaa tacgcaaacc gcctctcccc 5520  
gcgcgttgcg cgttgcattt atgcagctgg caacgacagg ttcccgactg gaaagegggc 5580  
agttagcgca acgcaattaa tggatgttag ctcaacttgg aggcacccca ggctttacac 5640  
tttatgtttc cggctcgat gtgtgtgaa attgttagcg gataacaatt tcacacaggaa 5700  
aacagctatg accatgatta cgaattgcgg ccg 5733

<210> 13  
<211> 39  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: PCR 5' primer  
  
<400> 13  
gtcaaccgcg gactgcgcattt catgaagttt ttcggccatc 39  
  
<210> 14  
<211> 66  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: PCR 5' primer  
  
<400> 14  
tcttagcaagc ttggctcttag ttctggaacc gcaccaggcg gcagcaacgg caacggcaat 60  
gtttgc 66  
  
<210> 15  
<211> 30  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: PCR 3' primer  
  
<400> 15  
tcgtacaagc tttcaagcac cgacggcggt 30  
  
<210> 16  
<211> 33  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: PCR 5' primer  
  
<400> 16  
tcttagctcta gaagcaacgg caacggcaat gtt 33

<210> 17  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 17  
tgcttagtcga cctgctagca gcaccgacgg cggtctg

37

<210> 18  
<211> 4614  
<212> DNA  
<213> *Saccharomyces cerevisiae*

<220>  
<223> *S. cerevisiae* FLO1 coding sequence

<400> 18  
atgacaatgc ctcatcgcta tatgtttttg gcaagtcttta cacttctggc actaaactagt 60  
gtggcctcag gagccacaga ggcgtgttta ccaggcggcc agaggaaaaag tgggatgaat 120  
ataaaatttt accatgttccat attgaaagat tcctccacat attcgaatgc agcatatatg 180  
gcttatggat atgcctcaaa aaccaaacta ggttctgtcg gaggacaaac tgatatctcg 240  
attgattata atattccctg tggtagttca tcaggcacat ttccctgtcc tcaagaagat 300  
tcctatggaa atggggatg caaaggaaatg ggtgcttgc ctaatagtca aggaattgca 360  
tactggagta ctgatttatt tggttctat actaccccaa caaacgttaac cctagaaaatg 420  
acaggttatt ttttaccacc acagacgggt tcttacacat tcaagttgc tacagttgac 480  
gactctgcaa ttctatcagt agggttgtca accgcgttca actgttgtgc tcaacagcaa 540  
ccgcccgtatca catcaacgaa ctttaccatt gacggatatac agccatgggg tggaaagttt 600  
ccacctaata tcgaaggaac cgtctatatg tacgctgttgc actattatcc aatgaaggtt 660  
gtttactcga acgctgttgc ttggggtaca cttccaattt gtgtgacact tccagatgg 720  
accactgtaa gtgatgactt cgaagggtac gtctattctt ttgacgtga cctaaagtcaa 780  
tctaactgta ctgtccctga cccttcaaat tatgctgtca gtaccactac aactacaacg 840  
gaaccatgga cccgtacttt cacttctaca tctactgaaa tgaccacgt caccggtaacc 900  
aacggcggttc caactgacga aaccgtcatt gtcatcagaa ctccaacaac tgcttagcacc 960  
atcataacta caactgagcc atgaaacagc acttttacact ctacttctac cgaatttgacc 1020  
acagtctgtc gcaccaatgg tgcacactt gacgaaacca tcattgtat cagaacacca 1080  
acaacagccca ctactgccccat aactacaactt gacccatgga acagcacttt tacctctact 1140  
tctaccgaaat tgaccacagt caccggtaacc aatggtttgc caactgtatga gaccatcatt 1200  
gtcatcagaa caccaacaac agccactact gccatgacta caactcagcc atgaaacgac 1260  
acttttacactt ctacatccac tgaatgacc accgtcaccc gtaaccaacgg tttgccaact 1320  
gatgaaacca tcattgtcat cagaacacca acaacagccca ctactgctat gactacaact 1380  
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccacgt caccggtaacc 1440  
aacggtttgc caactgatga aaccatcatt gtcatcagaa caccaacaac agccactact 1500  
gccatgacta caactcagcc atgaaacgac acttttacactt ctacatccac tgaatgacc 1560  
acggtcaccc gtaaccaatgg tttgccaact gatgagacca tcattgtcat cagaacacca 1620  
acaacagccca ctactgccccat gactacaactt cagccatgga acgacacttt tacctctaca 1680  
tccactgaaa tgaccacgt caccggtaacc aacggtttgc caactgtatga aaccatcatt 1740  
gtcatcagaa caccaacaac agccactactt gccataacta caactgagcc atgaaacgac 1800  
acttttacactt ctacttctac cgaatttgacc acagtccaccc gtaaccaatgg tttgccaact 1860  
gatgagacca tcattgtcat cagaacacca acaacagccca ctactgccccat gactacaact 1920  
cagccatgga acgacacttt tacctctaca tccactgaaa tgaccacgt caccggtaacc 1980  
aacggtttgc caactgatga aaccatcatt gtcatcagaa caccaacaac agccactact 2040  
gccatgacta caactcagcc atgaaacgac acttttacactt ctacatccac tgaatgacc 2100  
acggtcaccc gtaaccaacgg tttgccaact gatgagacca tcattgtcat cagaacacca 2160

acaacagcca ctactgccat gactacaact cagccatgga acgacacttt tacctctaca 2220  
tccactgaaa tgaccaccgt caccggtaacc aacggcggttc caactgacga aaccgtcatt 2280  
gtcatcagaa ctccaaactag tgaagggtcta atcagcacca ccactgaacc atggactgg 2340  
actttcacct ctacatccac tgagatgacc accgtcacccg gtactaacgg tcaaccact 2400  
gacgaaaccg tgatgttat cagaactcca accagtgaag gtttggttac aaccaccact 2460  
gaaccatgga ctggtaactt tacccatcata tctactgaaa tgaccaccat tactggaacc 2520  
aacggcggttc caactgacga aaccgtcatt gtcatcagaa ctccaaaccag tgaagggtcta 2580  
atcagcacca ccactgaacc atggactgg acttttactt ctacatctac taaaatgacc 2640  
accattactg gaaccaatgg tcaaccaact gacgaaaccg ttattgttat cagaactcca 2700  
actagtgaag gtctaatcag caccaccact gaaccatgga ctggtaactt cacttctaca 2760  
tctactgaaa tgaccaccgt caccggtaacc aacggcggttc caactgacga aaccgtcatt 2820  
gtcatcagaa ctccaaaccag tgaagggtcta atcagcacca ccactgaacc atggactggc 2880  
actttcacct cgacttccac tgaggttacc accatcactg gaaccaacgg tcaaccaact 2940  
gacgaaaccg tgatgttat cagaactcca accagtgaag gtctaatcag caccaccact 3000  
gaaccatgga ctggtaactt cacttctaca tctgtgaaa tgaccaccgt caccggtaact 3060  
aacggtaac caactgacga aaccgtgatt gttatcagaa ctccaaaccag tgaagggttg 3120  
gttacaacca ccactgaacc atggactgg acttttactt cgacttccac taaaatgtct 3180  
actgtcaactg gaaccaatgg ctgccaact gatgaaactg tcattgttgt caaaactcca 3240  
actactgcca tctcatccag ttgtcatca tcatcttcag gacaaatcac cagctctatc 3300  
acgtcttcgc gtccaaattat tacccttattc tacccttagca atgaaacttc tgtgatttct 3360  
tccctcgtaa ttcttccttc agtcaacttct tctctatc tcccttcgtc agtcaatttct 3420  
tccctcgtca ttcttccttc tacaacaacc tccacttctat tattttctga atcatctaaa 3480  
tcatccgtca ttccaaaccag tagttccacc tctggttctt ctgagagcga aacgagttca 3540  
gctggttctg tcttttcttc ctcttttattc tcttctgaat catcaaatac tccctacatat 3600  
tcttcttcat cattaccact tgttaccgt ggcacaacaa ggcaggaaac tgcttcttca 3660  
ttaccacctg ctaccactac aaaaacgagc gaacaaacca ctgggttac cgtgacatcc 3720  
tgcgagtc tctgtgcac tgaatccatc tcccttcgtca ttgttccac agctactgtt 3780  
actgttagc gcgtcacaac agagtatacc acatggtgcctt ctatttctac tacagagaca 3840  
acaagaaaa ccaaagggac aacagagcaa acccagaaaa caacaaaaca aaccacggta 3900  
gttacaattt ctcttcttca atctgacgtt tgctctaaga ctgcttcttc agccattgtt 3960  
tctacaagca ctgctactat taacggcggtt actacagaat acacaacatg gtgtcctatt 4020  
tccaccacag aatcgaggca acaaacaacg ctatgtactg ttacttctt cgaatctgg 4080  
gtgtgttccg aaactgcttc acctgcccatt gtttcgacgg ccacggctac tttgtatgt 4140  
gttggtaacgg tctatcctac atggaggccaa cagactgca atgaagatc tttgtatgt 4200  
aaaatgaaca gtgttccgg tgagacaaca accaataactt tagctgtca aacgactacc 4260  
aataactgttag ctgctgagac gattaccaat actggagctg ctgagacgaa aacagtagtc 4320  
accttcttcgc ttcaagatc taatcacgtt gaaacacaga cggctccgc gaccgatgtg 4380  
attggtcaca gcagtagtgtt tgtttcttca tccgaaactg gcaacaccaa gagtctaaca 4440  
agttccgggt tgagtagtat gtgcacacg cctcgttagca caccacgaa cagcatggta 4500  
ggatatagtt cagtttctt agaaatttca acgtatgtt gcagtgccaa cagttactg 4560  
gccggtagt gtttaagtgtt ctccattgtcg tttttatgtt ttaa 4614

<210> 19  
<211> 5857  
<212> DNA  
<213> *Saccharomyces cerevisiae* and *E. coli*

<220>  
<221> promoter  
<222> (1)..(452)  
<223> *S. cerevisiae* GAL1 promoter

<220>  
<223> (476-495) *E. coli* T7 promoter/priming site

<220>  
<223> (502-601) *E. coli* multiple cloning site

<220>  
<223> (609-857) *S. cerevisiae CYC1* transcription terminator

<220>  
<223> (1039-1712) *E. coli pMB1* (pUC-derived) origin

<220>  
<221> gene  
<222> (1857)..(2717)  
<223> *E. coli* ampicillin resistance gene

<220>  
<221> gene  
<222> (2735)..(3842)  
<223> *S. cerevisiae URA3* gene

<220>  
<223> (3846-5317) *S. cerevisiae* 2 micron origin

<220>  
<223> (5385-5840) *E. coli* f1 origin

<220>  
<223> (1-5857) Sequence of pYES2

<400> 19

acggattaga agccgcccag cgggtgacag ccctccgaag gaagactctc ctccgtcggt 60  
cctcgccctc accggtcgcg ttctgaaac gcagatgtgc ctgcgcgcg actgctccga 120  
acaataaaga ttctacaata ctatcttttta tggttatgaa gaggaaaaat tggcagtaac 180  
ctggccccac aaaccttcaa atgaacgaat caaattaaca accataggat gataatgcga 240  
ttatgtttttt agccttattt ctgggttaat taatcagcga agcgtatgtt tttgtatctat 300  
taacagatata ataaaatgcaa aaactgcatt aaccacttta actaataactt tcaacattti 360  
cggtttgtat tacttcttat tcaaataatgtaa taaaatgtatc aacaaaaat tggtaatata 420  
cctctataact ttaacgtcaa ggagaaaaaa ccccgatcg gactactagc agctgtataa 480  
cgactcaacta tagggaaat taagtcttgtt accgagatcg gatccactag taacggccgc 540  
cagtgtgctg gaattctgcgat gatatccatc acactggcgg cgctcgagc atgcatctag 600  
aggggcccat catgtatattt gttatgtcact gttatcgatc acggccatccc cccacatccg 660  
ctctaacgcgaa aaaggaagga gtttagacaat ctgaaatgtcta ggtccctatt tattttttta 720  
tagttatgtt agtattaaaga acgttatttttatttcaat ttttttttttttttctgtaca 780  
gacgcgtgtaa cgcgtatgaa attataactgaa aaaccttgct tgagaaggtt ttgggacgct 840  
cgaaggctt aatttgcggc cctgcattaa tgaatcgccc aacgcgcggg gagaggcgg 900  
ttgcgttattt ggcgttcccg cgcttcctcg ctcaactgact cgctgcgtc ggtcggttgg 960  
ctgcggcggag cggatcatcgc tcactcaaag gcggtatcac ggttatccac agaatcagg 1020  
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa agcccaggaa ccgtaaaaag 1080  
ggccgcgttgc tggcgaaaaa ccataaggctc cgccccctgt acgagcatca caaaaatcga 1140  
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gttttccct 1200  
ggaagctccc tcgtgcgtc tccgttcccg accctgcgcg ttaccggata cctgtccgc 1260  
tttctccctt cgggaagctgtt ggcgtttctt catagctcac gctgttagta tctcagttcg 1320  
gtgttaggtcg ttcgctccaa gctgggtctgt gtgcacgaaac ccccggtca gcccggaccgc 1380  
tgcgccttat cccgttaacta tcgtcttgag tccaaacccgg taagacacga cttatcgcca 1440  
ctggcagcag ccactggtaa caggattagc agagcgaggat atgttagggcg tgctacagag 1500  
ttcttgaagt ggtggctaa ctacggctac actagaagggc cagtttttttgcgatcgct 1560  
ctgctgtggc cagttacccctt cggaaaaaaga gttggtagctt cttgtatccgg caaacaacc 1620  
accggctggta ggggtgggtt ttttttttgc aaggcggaga ttacgcgcag aaaaaaaaaaagga 1680  
tctcaagaag atcccttgcgat cttttctacg ggggtctgacg ctcagtgaa cggaaaactca 1740  
cggttaaggga tttttggcat gagattatca aaaaggatct tcaccttagat ccttttaaat 1800  
taaaaaatgaa gttttaaatc aatctaaatgat atatatgat aaacttggtc tgacagttac 1860  
caatgcttaa tcgtgaggc acctatctca gcgatctgc tatttcgttc atccatagtt 1920

gcctgactcc ccgtcggtga gataactacg atacgggagc gcttaccatc tggccccagt 1980  
gctgcaatga taccgcgaga cccacgctca cccgctccag atttatcagc aataaaaccag 2040  
ccagccggaa gggccgagcg cagaagtggt cctgcaacctt tatccgcctc cattcagtt 2100  
attaattgtt gccgggaagc tagagtaatg agtgcgcag ttaatagttt gcgaacgtt 2160  
gttggcattt ctacaggcat cgtgggtgtca ctctcgctgt ttggatggc ttcatcagc 2220  
tccgggttccc aacgatcaag gcgagttaca tgatccccaa tggtgtgcaa aaaagcggtt 2280  
agtccttcg gtcctccgt cgttgcaga agtaagttgg ccgcagttt atcactcatg 2340  
gttatggcag cactgcataa ttcttactt gtcatgccat ccgtaaatgtt cttttctgt 2400  
actggtgatg actcaaccaa gtcatctgtaa gaatagtgtt tgccgcacc gagttgtct 2460  
tgcccgccgt caatacggtt taatagtgtt tcacatagca gaattttaaatgtt 2520  
attggaaaac gttttcggtt gcgaaaaactc tcaaggatctt taccgtgtt gagatccagt 2580  
tcgatgtaac ccactcgatc acccaactga ttttcagcat cttttactttt caccagcggtt 2640  
tctgggttagt caaaaacagg aaggcaaaat gcccggaaaaa aggaaataag ggcgacacgg 2700  
aaatgttcaa tactcataact cttttttttt caatgggtaa taactgatataattttt 2760  
aagctctaattt ttgtgatgtt agtatacatg cattttacttta taatacagttt tttttttttt 2820  
gctggccgca tcttctcaaa tatgttttttcc accgtgtttt tctgttaacgt tcaccctcta 2880  
cctttagcattt cttttttttt gcaaatagtc ctcttccaaac aataataatgt tcaagatctt 2940  
tagagaccac atcatccacg gttttttttt gttttttttt tgccgtctccc ttgttcatctt 3000  
aaccccacacc ggggtgttata atcaaccaat cgtaacccctt atcttcttcca cccatgtctt 3060  
ttttagtataat aaagccgatata acaaaatctt tgccgtctt ccgtaaatgtt acagtaccctt 3120  
tagtatattt tccagtagat aggaggccctt tgcatgacaa ttctgttaacatc atcaaaaggc 3180  
ctcttaggtttt cttttttttt cttttttttt cttttttttt accgttttttccaa ataccttggc 3240  
ccaccacacc gtgtgttccatc gtaatgttccatc tatttttttttccaa acacccgcag 3300  
agtactgcaat tttttagtataat ttaccaatgtt cggaaatttt tctgttttttccaa aagagttttttt 3360  
aattttttttt ggcggataat gcttttttttccaa gctttaactgtt gcccctccatg gaaaaatccatg 3420  
tcaagatataatc cccatgtttaat tttagtataat tttagtataat tttagtataat tttagtataat 3480  
ccagtttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 3540  
gcatgatattt aaatagtttgc gcaacatccatc atgttttttttccaa ctttttttttccaa 3600  
atgttagttt ccacatgtttaat tttagtataat tttagtataat tttagtataat tttagtataat 3660  
ttaagaataatc tggggcaattttt catgttttttccaa ctttttttttccaa ctttttttttccaa 3720  
taagtctgtt ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 3780  
tttcaagaaaaa ccgaaatccatc aaaaaaagaat aaaaaaaaaaa tgatgttataat tttagtataat 3840  
cttagtttttttccaa gatgttttttttccaa gatgttttttttccaa gatgttttttttccaa 3900  
cttagtttttttccaa gatgttttttttccaa gatgttttttttccaa gatgttttttttccaa 3960  
cttaccacccatc ttttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4020  
ctatcttcgc gatgttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4080  
acttttttttccaa tgggtgttccatc ttttttttttccaa ctttttttttccaa ctttttttttccaa 4140  
tcgatgttttttccaa ttttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4200  
tcgtacttgtt ttttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4260  
agatgttataat ttgttcaatgtt ttttttttttccaa ctttttttttccaa ctttttttttccaa 4320  
atgttatttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4380  
atccccgggtt ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4440  
agcatctgtt ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4500  
aaagaatctt ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4560  
acgaagaatccatc ttttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4620  
tttcaaaacaa agaatctgtt ctttttttttccaa ctttttttttccaa ctttttttttccaa 4680  
tttttcaaaacaa agaatctgtt ctttttttttccaa ctttttttttccaa ctttttttttccaa 4740  
tatttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4800  
gttcttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4860  
tcttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4920  
gaagctgttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 4980  
ggatgtgttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5040  
aattatgttccatc ttttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5100  
ttcgtatttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5160  
gtaatactgtt ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5220  
cgaaagggttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5280  
cttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa ctttttttttccaa 5340  
tcagaaaaacatc cccaaaaacatc ggaagattgtt ataaacgttttccaa ctttttttttccaa 5400

tatTTTgtta aaattcgcgt taaatTTTtg ttAAATcAGC tcATTTTta acGAATAGCC 5460  
cgAAATCGGc AAAATCCCTt atAAATCAAa agAAATAGACC gagATAGGGt tgAGTGTtGt 5520  
tCCAGTTCC aacaAGAGTC cactATTAAa gaACGTGGAC tCCAACGTCA aAGGGCGAAA 5580  
aAGGGTCTat cAGGGCGATG gCCCACtACg tGAACCATCA ccCTTAATCAA gTTTTTGGG 5640  
gtcGAGGTGc CGTAAAGCAG tAAATCGGAA gGGTAAACGG ATGCCCCAT tTAGAGCTG 5700  
acGGGGAAAG CGGGCGAACG TGGCGAGAAA gGAAGGGAAg AAAGCGAAAAG gAGCGGGGGC 5760  
tagggcggtg ggaAGTGTAG gggTCACGCT gggCGTAACC accACACCCG CCGCGCTAA 5820  
tggggcgcta cagggcgctg ggggatgata cactagt 5857

<210> 20  
<211> 403  
<212> DNA  
<213> Trichoderma reesei

<220>  
<223> (1-403) T. reesei hfb2 coding sequence

<220>  
<221> intron  
<222> (131)...(200)

<220>  
<221> intron  
<222> (287)...(358)

<400> 20  
atgcagttct tcGCCGTCGC CCTTTGCCC accAGCGCCC tggCTGTGT ctGCCCTacc 60  
ggcCTCTCT ccaACCCtCT gtGCTGTGCC accAAACGTCC tcGACCTCAT tggCGTTGAC 120  
tgcaAGACCC gtATGTTGAA ttccaATCTC tggcATCTC gacATTGGAC gatacAGTTG 180  
acttACACGA tgCTTTACAG ctACCATCGC cgtcGACACT ggCCCATCT tCCAGGCTCA 240  
ctgtGCCAGC aAGGGCTCCA agcCTTTG ctgcgttGt cccgtggtaa gtagtGCTG 300  
caatggcaAAa gaAGTAAAAA gacATTGGG cctgggatcg ctaACTCTG atATCAAGGC 360  
cgaccAGGCT ctccTGTGCC agaAGGCCAT cggCACCTC taa 403

<210> 21  
<211> 59  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 21  
cgaggAGGCT cgacGACTTC gAGCAGCCCG agctgcACGC aggCTGTCTG ccCTACCGG 59

<210> 22  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 22  
tcattggatc cttagaAGGT gcccGATggc

<210> 23  
<211> 679  
<212> DNA  
<213> *Schizophyllum commune*

<220>  
<223> (1-679) SC3 coding sequence

<220>  
<223> (1-92) 1st cDNA

<220>  
<223> (146-183) 2nd cDNA

<220>  
<223> (240-317) 3rd cDNA

<220>  
<223> (374-469) 4th cDNA

<220>  
<223> (524-586) 5th cDNA

<220>  
<223> (635-679) 6th cDNA

<400> 23  
atgttcgccc gtctccccgt cgtgttcctc tacgccttcg tcgcgttcgg cgccctcg 60  
gctgcccctcc cagggtggcca cccgggcacg acgtacgtcg acctctcacc gtccctcta 120  
gtcttgcgtga tgaagccccg tatagcacgc cgccgggtac gacgacggtg acggtgacca 180  
cggtgagtag ctttctcgcc gtcgacgact cgaacgcatt ggctaatttt tgctctatgc 240  
cgccctcgac gacgaccatc gecgcgggtg gcacgtgtac tacggggtcg ctctcttgct 300  
gcaaccaggta tcaatcggtt cgtacatcaa agcgggcacga ccaggcatct cagctgacgg 360  
ccacatcgta caggcgagca gcagccccgt taccggccctc ctcggccctgc tcggcattgt 420  
cctcagcgcac ctcaacgttc tctgtggcat cagctgcctc cccctcaactg tgagatctt 480  
ttgttcaactg tcccaattac tgcgcactga cagacttgc caggtcatacg gtgtcgagg 540  
cagcggtgtt tcggcgcaga ccgtctgtg cggaaacacc caattcgat gtatacttgc 600  
catgcgtgtc cttttctccg ctaatcatct gtagaacggg ctgatcaaca tcggttgcac 660  
ccccatcaac atcctctga 679

<210> 24  
<211> 63  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 24  
actacacgga ggagctcgac gacttcgagc agcccgagct gcacgcaggg tggccacccg 60  
ggc 63

<210> 25  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 25  
tcgtacggat cctcagagga tggatggg 30

<210> 26  
<211> 43  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 26  
gaaattccgc ggactgcgca tcatacgatt ctgcgcattt gcc 43

<210> 27  
<211> 80  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 27  
tgaattccat atgttaggtt ccacccgggc ccatgcgggt agaagttagaa gccccgggag 60  
caccgacggc ggtctggcac 80

<210> 28  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 28  
tgaattcgggt acccaggctt gctcaagcgt c 31

<210> 29  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 29  
tgaattccat atgtcacagg cactgagagt agta 34

<210> 30  
<211> 48  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 30  
gaattcggta ccctcgccc tcgcggtccc gccgaagtga acctggtg 48

<210> 31  
<211> 34  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 31  
tgaattccat atgctaacc cgtttcatct ccag 34

<210> 32  
<211> 918  
<212> DNA  
<213> Trichoderma reesei

<220>  
<221> terminator  
<222> (1)..(918)  
<223> T. reesei hfbl terminator

<400> 32  
gatccccgcc cggggtaaag gtgtgcccgt gagaaaagccc acaaagtgtt gatgaggacc 60  
atttccggta ctggaaagt tggctccacg tgtttggca gggttggca agtttgttag 120  
atattccatt cgtacgcacat tcttattctc caatatttca gtacacttt cttcataaat 180  
caaaaaagact gctattctt ttgtgacatg cccgaaggga acaattgttc ttggctcttg 240  
ttatttgcaa ttaggatgg gagattcgcc tttagagaaag tagagaagct gtgcttgacc 300  
gtggtgtgac tcgacgagga tggactgaga gtgttaggt taggtcaac gttgaagtgt 360  
atacaggatc gtctggcaac ccacggatcc tatgacttga tgcaatgggt aagatgaatg 420  
acagtgttaag aggaaaagga aatgtccgcc tttagctgtat atccacgccca atgatacagc 480  
gatataccctc caatatctgt gggAACgaga catgacatcat ttgtggaaac aacttcaaac 540  
agcgagccaa gacctcaata tgcacatcca aagccaaaca ttggcaagac gagagacagt 600  
cacattgtcg tcgaaagatg gcatctgtacc caaatcatca gctctcatca tcgcctaaac 660  
cacagattgtt ttgcgcgtccc ccaactccaa aacgttacta caaaagacat gggcaatgc 720  
aaagacccatga aagcaaaccc ttttgcgac tcaattccct cctttgtcct cggaatgtatg 780  
atcccttcacc aagtaaaaga aaaagaagat tgagataata catgaaaagc acaacggaaa 840  
cgaaaagaacc aggaaaagaa taaatctatc acccaccttg tccccacact aaaagcaaca 900  
gggggggtaa aatgaaat 918

<210> 33

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 5' primer

<400> 33

gacctcgatg cccggccggg gtcaag

26

<210> 34

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR 3' primer

<400> 34

gtcgacattt cattttaccc ccctcg

26

<210> 35

<211> 1190

<212> DNA

<213> Trichoderma reesei

<220>

<221> promoter

<222> (1)..(1190)

<223> T. reesei hfb2 promoter

<400> 35

ctcgaggcgc tgaagcttgc atgcctgcat cttttgttag cgactgcac cattttgcac 60  
acactgcgt cgacgtctct ctccgcacct tggccagctg gacaagcaac acaccaatga 120  
cgctttgtat tattagatgtat tatgcaagtc tcaggactat cgactcaact ctacccaccg 180  
aggacgatcg cggcacata cgccctcggt ctcattggcc caagcagacc aactgcccct 240  
ggagacaatgc tcagccaaag ggagatggac ggcaggggcac gccaggcccc caccaccaag 300  
ccactccctt tggccaaatc agtttgcatt tcaagagaca tcgagctgtg ccttgaaatt 360  
actaacaacc agggatggaa aacgaaggct gcttttggaa agacaacaat gagagagaga 420  
gagagagggaa gagagacaat gagtgccaca aacctggtag tgctccgcca atgcgtctga 480  
aatgtcacat ccgagctttg gggcctctgt gagaatgtcc agagtaatac gtgtttgcg 540  
aatagtcctc ttctttgttggagg actggatacc tacgataacc tttttggatgtt gatgcgggtgc 600  
tttcgttggatgtt ttatctggag gatagaagac gtcttaggtaa ctacacaaaa ggcctataact 660  
tttggggatgtt gcccaacgaa aggttaactcc tacggccctt tagagccgtc atagatccca 720  
cagccctttg gagccgtcat agatcacatc tggtagagacc gacattctat gaataatcat 780  
ctcatcatgg ccacatacta ctacatacgat gtctctgcct acctgacatg tagcagtggc 840  
caagacacca aggccccagc atcaagccctc cttttccatgt tacagcggca 900  
gagagattgc gatgaggccctt ctcccttaccc acagacggct gacaatgtcc gtataaccacc 960  
agccaaacgtt atgaaaacaa ggacatgagg aacagccctgc gagagctgaa agatgaagag 1020  
ggcccgaaaaaaa aaaagtataa agaagacctc gattcccgcc atccaacaaat ctttccatc 1080  
ctcatcatgg cactcatcta caaccatcac cacattcaactt caactccctt ttctcaactc 1140  
tccaaacaca aacattttt gttgaatacc aaccatcacc acctttcaag 1190

<210> 36  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 5' primer

<400> 36  
aagcttgcatttgcattcc

20

<210> 37  
<211> 26  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: PCR 3' primer

<400> 37  
ccatggtgaa aggtgggtat ggttgg

26

<210> 38  
<211> 13  
<212> PRT  
<213> Trichoderma reesei

<220>  
<223> wild type *T. reesei* EGI peptide linker

<400> 38  
Val Pro Arg Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly  
1 5 10

<210> 39  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: modified CBHII linker

<400> 39  
Gly Ser Ser Ser Gly Thr Ala Pro Gly Gly  
1 5 10

<210> 40  
<211> 19  
<212> PRT  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Met /Thrombin linker

&lt;400&gt; 40

Pro Gly Arg Pro Val Leu Thr Gly Pro Gly Met Gly Thr Ser Thr Ser  
1 5 10 15

Ala Gly Pro

&lt;210&gt; 41

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Met-containing linker

&lt;400&gt; 41

Pro Gly Ala Ser Thr Ser Thr Gly Met Gly Pro Gly Gly  
1 5 10

&lt;210&gt; 42

&lt;211&gt; 14

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: linker containing the thrombin cleavage site

&lt;400&gt; 42

Gly Thr Leu Val Pro Arg Gly Pro Ala Gly Val Asn Leu Val  
1 5 10